

CHAPTER 12

Poverty-Sensitive Scorecards to Prioritize Lending and Grant Allocation with an Application in Central America

Manuel A. Hernandez¹ and Máximo Torero²

1 Introduction

The importance of credit access in improving economic opportunities in developing markets is well established and generally recognized by policymakers. The optimal use of loan funds, however, could be subject to a potential tradeoff between sustainability and poverty reduction. This chapter develops a poverty-sensitive scorecard system for lending or grant allocation in underdeveloped markets. The methodology is innovative in that it combines both a risk and a poverty scorecard. This ensures that the loan or grant allocation is not only focused on targeting the poor, but also on ensuring the sustainability of the investment project. In addition, the risk scoring uses an innovative non-linear and nonparametric model that leads to better assessment of credit worthiness and results in a lower screening of the poor from this extremely important market. We then implement the scorecard system using a real example of grant competition in Central America to link rural smallholders to markets.

The importance of credit in improving economic opportunities in developing markets is well documented (Armendariz de Aghion and Morduch 2005; Shahidur, 2006; Brett, 2006, Gosh, Mookherjee and Ray, 2000). Similarly, the role of micro- and small enterprises in economic development has been recognized by policymakers since the 1970s when the overwhelming patronage of large-scale industries did not bring redistribution with growth (McPherson, 1996). Promotion of small enterprises requires a leveling of the playing field between the large-scale sector and the small-scale sector. One way to do so is to design policies that do not explicitly or implicitly discriminate against the small-scale sector. Policies that assist small enterprises may also be desirable. Measures to assist small enterprises

¹ Postdoctoral Fellow at Markets, Trade, and Institutions Division, IFPRI,
e-mail: m.a.hernandez@cgiar.org.

² Director of Markets, Trade, and Institutions Division, IFPRI,
e-mail: m.torero@cgiar.org. All inquires should be sent to Maximo Torero.

have included training in production and marketing methods, as well as measures to make credit available to them.³ Apart from priority and subsidized lending, financial assistance can be provided by donors in the form of grants.

In lending or extending grants to small enterprises, the choice of projects among competing demanders is important. In underdeveloped markets where contracts are difficult to enforce and the problem of adverse selection (wrong choices made when the type of borrower is unknown) is severe, lending risks are high. Traditionally, credit in these sectors has been characterized by money lenders offering loans with relatively high interest rates. The credit from government banks to the majority of poor borrowers has been restricted due to stringent collateral demands. One form in which formal credit to the poor has expanded is through microfinance. This type of lending is targeted to self-employment activities without accompanying collateral. It is well known that the absence of collaterals can lead to credit market failure (Ghatak, and W. Guinnane, 1999) when a loan is granted to individuals rather than groups.

Most of microfinance credit is extended based on the reputation of the borrower (for example his/her past borrowing and repayment pattern). As reputation is difficult to measure, the lack of a more precise measure of borrower's riskiness affects the efficiency of the credit market. Thus, in spite of the well-publicized success stories in microfinance credit, there have been concerns that the lending institutions have been able to sustain low rates of interest and relatively high default rates mainly due to subsidies and soft loans. For example, Grameen Bank charges an average real rate of 10 percent, and has experienced losses close to 18 percent of outstanding loans over the period 1985–1996 if proper adjustment for the portfolio size is taken into account (Armendaris de Aghion and Morduch, 2005).

Banks face greater difficulty in overcoming adverse selection problems in individual lending. A credit scorecard that predicts credit worthiness of the borrower accurately can address the problem of market failure to a large extent. With a credit score, lending institutions will also be able to offer a menu of choices combining interest rates and approved loan amount for a potential borrower. The fundamental element in producing such a menu is the creation of an accurate risk ranking (hitherto missing in case of most lending to the poor including in microfinance credit) for individual borrower units based on some specific attributes. Credit scoring models have been shown to be the most effective tool in determining the riskiness of a borrower and are usually based on long historical data involving several entities (for example, an integrated system of lending institutions).

However, the riskiness of a borrower (in terms of the chance of default on repayment for loans and in terms of the efficient and adequate use of funds for grants) is not the only criterion that should be used in the case of development

³ The vast majority of these credit programs, especially the so-called "agricultural development banks", which provided credit at subsidized interest rates, have failed to achieve their objectives both to serve the rural poor and be sustainable credit institutions (Adams, Graham, and von Pischke, 1984; Braverman and Guasch, 1986; Adams and Vogel, 1986).

lending. The menu of projects should also be assessed in terms of their potential for reducing poverty. Consequently, the optimal use of loan funds could be subject to a potential tradeoff between sustainability and poverty reduction.

This chapter develops a poverty-sensitive scorecard system for lending or grant allocation in underdeveloped markets, which addresses the question of how to lend or provide grants when faced with the above-mentioned tradeoff. The proposed methodology consists of two stages. The first accounts for the development of a risk scorecard – a risk score – for potential borrowers using the latest developments in econometric modeling. The nonparametric technique we propose will allow risk ranking through credit scores that significantly improve upon current methods by providing a more accurate measure of risk associated with individual loans, and more importantly, provide a more accurate measurement of potential gains and losses associated with each loan. In the second stage, following a Principal Components approach, a poverty scorecard is developed to evaluate the projects of potential borrowers who have a risk score below a certain threshold. That is, projects that are proven to be sustainable in the first stage are ranked in the second stage in terms of their potential impact on poverty reduction.

The remainder of the chapter is organized as follows. Section 2 discusses the proposed risk and poverty scorecards in detail, including the methodology used to develop and implement them. In Section 3, we apply the poverty scorecard system to a real example of grant competition in Central America to link rural small-holders to markets. Section 4 provides concluding remarks.

2 Building a Scorecard

The scorecard system proposed for lending or grant allocation (extension) in developing markets consists of two stages. In the first stage, a risk scorecard or algorithm suitable for lending to small enterprises is constructed and applied using the latest developments in statistical (econometric) modeling. In the second stage, the projects of potential borrowers (beneficiaries) with a risk score below a particular threshold are evaluated in terms of their potential impact on poverty reduction through a poverty scorecard. The process involving the two stages is summarized in [Figures 1 and 2](#).

The conceptual framework behind the proposed risk and poverty scorecard system rests both on targeting the poor and on assuring sustainability of the project. With the premise that sustainability is a necessary condition for poverty reduction, in stage 1, a loan/grant application is evaluated in terms of the borrower's default probability. Similar to the default likelihood in lending markets, which is usually captured through a credit scorecard, a risk score or default probability can be computed using information from the loan/grant applications as well as information from other sources.

Note that the interpretation of default differs for a loan and a grant. In the case of a loan, default is to be interpreted as an event in which the loan is not repaid in time.

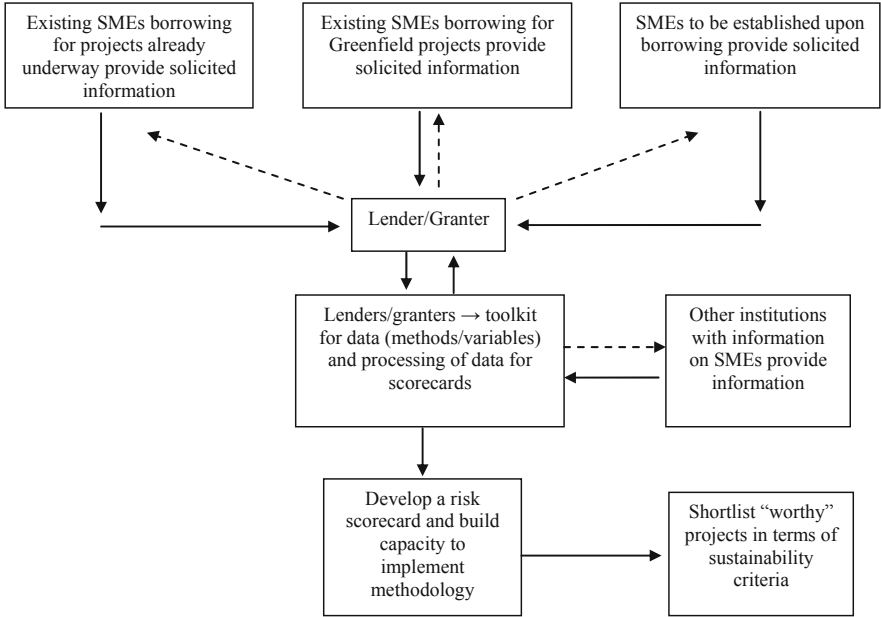


Fig. 1. Stage 1: Creating the Sustainability Cut-off Using a Risk Scorecard

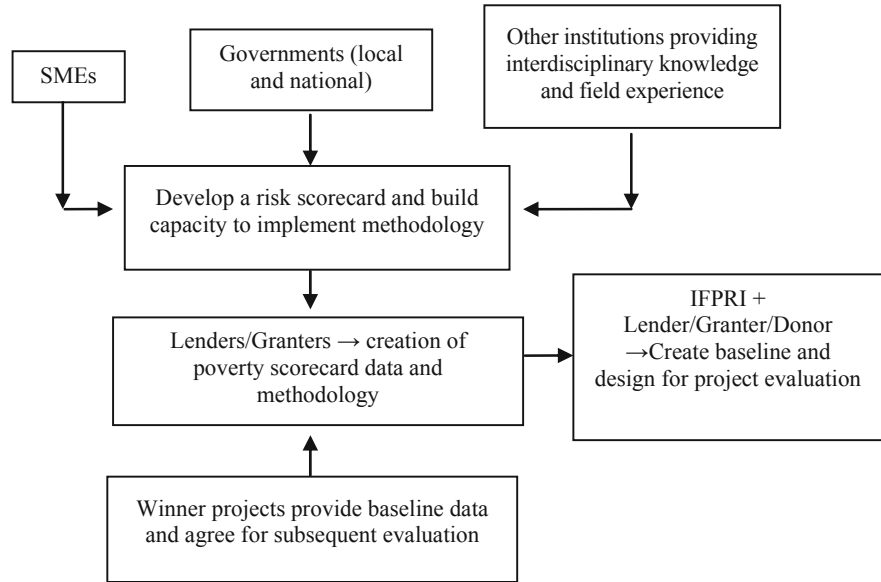


Fig. 2. Stage 2: Creating a Poverty Scorecard among Short-Listed Projects

In the case of a grant, where there is no repayment obligation, the sustainability of the project needs to be assessed in terms of alternative indicators. Thus, in the case of loans, a standard credit scorecard framework can be applied, while in the case of grants, estimating a survival probability or the probability of returns beyond a threshold could be assessed using data from a reference or a comparable group of enterprises.

Information from several sources can then be combined to estimate default probability for the loan/grant applications. Some data can be collected in conjunction with the loan/grant application. However, data regarding transactions of the applicant with other institutions, data regarding external factors that can affect the profitability/viability of the enterprise/project, and data from comparable enterprises are essential for improving the predictive power of the scoring model. After selecting those projects with a minimum risk level to assure their sustainability, these projects are then evaluated in terms of their potential impact on poverty.

It is worth noting that the creation of a scorecard system is a dynamic process that can be improved over time, particularly when there is initially limited information for risk assessment (see [Figure 3](#)). An initial risk algorithm can be derived based on the performance of comparable enterprises/borrowers. This algorithm can then be applied to evaluate the loan/grant applications of interest, using the information collected during the application process as well as information from other sources. The outcomes of the projects that were selected based on their risk score and potential impact on poverty reduction can then be further evaluated, which will help to expand and update the data for the scorecard system. Thus, an interim scorecard can be created (most likely through a pilot intervention) that not only exploits existing information but also incorporates additional incoming information to improve the scorecard. This process can continue until a performance history is established for a reasonably large portfolio of loans/grants.

Next, we discuss in more detail the methodology used to develop and implement the risk and poverty scorecards.

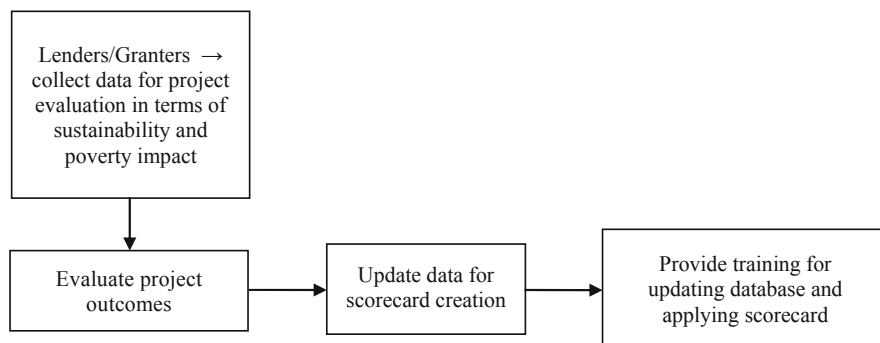


Fig. 3. Evaluation of Project and Strengthening of Database and Scorecard

2.1 Risk Scorecard

In well-developed credit markets, the use of credit scores is an integral part of the lending process. The increase in the use of credit scoring systems in developing economies, in contrast, is a relatively recent phenomenon. We intend to develop a credit scoring system for lending or extending grants to small enterprises that employs the latest developments in statistical (econometric) modeling, which are likely to improve the accuracy of risk ranking compared to existing methods. In particular, our objective is to develop a credit scoring algorithm suitable for lending to small and micro-enterprises and to provide the guidelines needed to generate a working database on credit scores based on available data and on potential data that can be collected in the future to improve the model prediction. It is also our intention to make the credit scoring mechanism simple to apply.

The development of a robust system of risk scoring will benefit both the lender/granter and the borrowers. The absence of a well-developed risk scoring system not only restricts access to credit but also prevents the development of differentiated borrowing options. Given a credit score, lending institutions are able to offer a menu of choices combining interest rates and (pre-approved) loan amounts for a potential borrower.⁴ The fundamental element in producing such a menu is the creation of an accurate risk ranking (which has been missing until now in the case of most lending to micro- and small enterprises), but the accuracy of this ranking mainly depends on the statistical model used to construct it. We propose a model that more accurately reflects the relationship between a borrower's characteristics and risk level.

As noted above, detailed information regarding both a borrower's specific characteristics and external factors plays a crucial role in building the risk-based scorecard. However, in developing countries, detailed information regarding these variables is generally not available; thus, loans/grants are sometimes granted by qualitative/subjective evaluation. The methodology proposed below is also applicable when there is initially limited information. New information can then be incorporated across time to improve the scorecard.

Methodology

Associated with every potential borrower, there is a probability of default conditional on the terms of the loan/grant being requested. This probability depends on a borrower's attributes as well as on external factors that are not borrower-specific. The main purpose of a scorecard is to create a ranking of borrowers by estimating these probabilities, recognizing that their magnitudes depend on several borrower attributes and factors. For example, consider a loan request from a small enterprise

⁴ Risk scoring is not necessarily an instrument used to discriminate against high risk, but an important tool in designing the portfolio of borrowing options. Hence, the term of loans/grants or the interest rates offered will tend to be sensitive to the type of borrower.

closely related to the agricultural sector. The borrower's characteristics may include current assets, income, credit history, and outstanding debt, while external factors may include crop price volatility. In general, estimation is conducted by relating a discrete (binary) variable to a borrower's characteristics and to external factors.

A suitable statistical model (regression) should be an accurate representation of the underlying relationship between attributes, external factors, and the defined binary variable. The statistical model used belongs to a class of discrete choice models in which the odds of a loan being of high (or low) risk is a function of borrower-specific characteristics and other attributes. The specific assumption about the functional form for this relationship becomes crucial for classifying the borrowers into risk categories. Depending upon whether the chosen functional form is correct or not, the accuracy of the model predictions is likely to vary significantly. It is also worth noting that functional assumptions imposed in the development of credit scores in developed countries is often of limited use in credit markets in developing countries, since key variables that affect the riskiness of borrowers, and the way in which they affect the level of riskiness, vary significantly between developed and developing economies.⁵ It is useful for developing countries to develop a methodology that does not depend on prior knowledge or particular assumptions on functional forms.

The specific methodological innovation that we propose consists of not assuming a specific functional and distributional form in the relationship between borrowers' characteristics, external factors, and the default probability. Allowing the data to fit in the best functional form is preferable to imposing specific (and most likely erroneous) functional and/or distributional assumptions. Statistical models in which specific functional forms are not imposed are known as semi-parametric and nonparametric estimation methods. Although very well suited for credit scoring research, these methods have surprisingly not been widely applied in this area. A plausible explanation for this lies in the fact that most implementable semi-parametric and nonparametric methods have been recently developed.

A simple example can illustrate the problem of assuming a specific functional form that happens to not be the correct one. Let the default probability of a borrower depend on the loan amount, debt ratio, and asset size. To demonstrate the utility of a data-driven method, assume that there exists an unknown threshold asset size below which the default probabilities rise exponentially. For asset sizes above the threshold level, the default probabilities do not depend on asset size. The standard scoring models assume that the odds of default are linear with respect to all explanatory variables and that the underlying distribution governing the relationship between the probability of default and the explanatory variables is known. The use of these methods will incorrectly estimate the risk of default for

⁵ For example, it is well established that the loan-to-collateral ratio and payment burden are significant contributing factors in risk assessment. However, in agriculture-based lending, price volatility and probability of crop failure will play a much bigger role in risk ranking.

borrowers with asset holdings below the threshold level. This may lead to the potential exclusion of “good” borrowers from the market.

The specific econometric method that we propose for estimating risk scores is the semi-parametric single index model derived by Klein and Spady (1993) because of its relatively faster and less computational burden estimation process, compared to a fully nonparametric method. Recall that we are also looking for a method that can be easily implemented. Compared to a parametric method, the semi-parametric single index model does not impose a specific distributional form when modeling the probability of default. In particular, the probability of default is given by,

$$P(Y = 1 | X) = E(Y | X) = g(X'\beta), \quad (1)$$

where Y is the binary variable associated with the default of past projects/loans, X is the set of socioeconomic and financial characteristics of the borrower, plus other factors that could affect the likelihood of default, $g(\cdot)$ is an unknown function, and β is the set of parameters to be estimated. This model is semi-parametric in nature since the functional form of the linear index is specified, while $g(\cdot)$ is left unspecified. In a parametric setting, $g(\cdot)$ could be assumed, for example, to be a Normal function, which would result in the widely used Probit model.

Klein and Spady suggest a semi-parametric likelihood approach to obtain parameters β from (1). More specifically, the maximum likelihood estimator of β is given by,

$$\hat{\beta} = \arg \max_{\beta} L_n(\beta, g) = \sum_{i=1}^n (y_i \ln g(X_i'\beta) + (1 - y_i) \ln(1 - g(X_i'\beta))), \quad (2)$$

where $g(\cdot)$ is approximated for each borrower/applicant i through a leave-one-out nonparametric kernel estimator, such that

$$\hat{g}_{-i}(X_i'\beta) = \frac{\sum_{j \neq i} k \left(\frac{(X_j - X_i)'\beta}{h} \right) y_j}{\sum_{j \neq i} k \left(\frac{(X_j - X_i)'\beta}{h} \right)}. \quad (3)$$

The risk score estimated for each potential borrower or grantee i is, then, the estimated \hat{g}_{-i} using $\hat{\beta}$.⁶

⁶ Single index models also require two identification conditions under which the unknown parameter vector β and the unknown function $g(\cdot)$ can be sensibly estimated.

Steps to Construct a Risk Scorecard

Given a pool of potential borrowers or grantees, the steps to derive a risk score can be summarized as follows:

1. Gather information from past loans (grants) from the same pool of borrowers (grantees) and/or from other comparable sets of loans (grants) from applicants similar to the pool of borrowers (grantees). The information collected should allow us to determine/identify the performance of those borrowers/grantees (default or sustainability of their project). The information should also include (at least) socioeconomic and financial characteristics of the applicants.
2. Estimate the performance of these previous borrowers/grantees as a function of their socioeconomic and financial characteristics and other related factors (if available) using the semi-parametric single index model described above.
3. The resulting estimated vector of parameters $\hat{\beta}$ from step 2 can, then, be used in (3) to derive a risk score for the interested pool of borrowers or grantees. We only need to provide the same set of characteristics and factors used in the estimation of β for the pool of potential borrowers/grantees to be evaluated.
4. Only those projects (borrowers/grantees) with a risk score below a certain threshold will then be assessed in terms of their potential impact on poverty.

Note that this risk scoring mechanism could operate through a simple implementable program in open-source software, such as a spreadsheet. The estimated vector of parameters $\hat{\beta}$ from step 2, and the procedure to obtain \hat{g}_{-i} , described in equation (3), can be embedded in the program; the user will then only need to provide the set of characteristics contained in X for the potential borrower/grantee(s) to be evaluated. However, given that there might initially be limited information regarding loans and grants, particularly for small enterprises and/or development projects, new information can be incorporated across time to improve the scorecard system, particularly the precision of the estimated $\hat{\beta}$ parameters. A pilot stage could be a plausible start so that an interim scorecard is built (with limited information) and then, based on the evaluated performance of the selected borrowers/grantees or projects, additional information can be subsequently incorporated

First, the set of explanatory variables X must contain at least one continuous variable. Second, β cannot be identified without some location and scale restrictions (normalizations). One popular location-normalization is to not include a constant in X ; one popular scale-normalization is to assume that the first component of X has a unit coefficient and that this first component is a continuous variable. For further details on semi-parametric single index models, refer to Li and Racine (2006).

to improve the scorecard. This process can continue until a performance history is well established for a reasonably large portfolio of loans/grants.

2.2 Poverty Scorecard

Conditional on meeting the sustainability cutoff, the final selection of projects requires assessing their poverty reduction potential. There are several indicators of the impact on poverty that could be employed. For example, projects/enterprises could be assessed in terms of their labor intensity (low-skilled, female labor) or in terms of their geographical location (by which those projects located in poorer areas are expected to have a greater impact on poverty).

Additionally, the impact of a project is likely to vary from one region to another depending upon local conditions and other factors, such as access to external markets and infrastructure. A project in areas with a high poverty rate and that face infrastructure constraints is expected to have a higher impact on the poor than a project in an area with fewer poor people and no binding constraints. Using poverty maps combined with market access information could be very useful in approximating a project's impact on poverty reduction. Maps of market access, defined in terms of estimated travel time to roads, markets, and/or cities, would be very helpful for sustainability analysis per se, but could also be used for poverty targeting.

Table 1. List of Potential Variables for the Creation of a Poverty Scorecard

Variables	Criteria for evaluation
Geographic indicators	
1. Location in poverty mapping space	The project should be executed in places with high poverty rates.
2. Access to markets	The project should be executed in places with low market access, e.g. very low accessibility to a main road or major market (city).
Employment indicators	
3. Labor intensity	Number of new jobs generated by the project.
4. Low-skill labor intensity	Percent of low-skill labor in the project.
5. Female labor intensity	Percent of female labor in the project.
Spillover indicators	
6. Effects on supply chain participants	Total number of direct beneficiaries from the project (ratio based on total amount invested).
7. Other spillover effects (indirect effects, provision of public good)	Total number of indirect beneficiaries from the project (ratio based on total amount invested).

A value chain perspective is also recommended when assessing a project's poverty impacts since there are possible spillover effects across the supply chain (for example, direct and indirect labor effects). The impact on vulnerable populations could also be considered (such as impacts on women's employment and children's education). A list of potential variables that could be included to develop a poverty scorecard is presented in [Table 1](#).

Methodology

When assessing a project's potential impact on the poor, complexity can arise in terms of weighting the outcomes. It is possible that some enterprises/projects perform better in terms of geographical targeting, but do not do well in terms of their potential impact on gender. Consequently, we propose the use of the statistical method of Principal Components to determine the weights for the different outcomes (variables) considered. Principal Component analysis is a statistical technique that creates new variables that are linear combinations of the original variables. The new variables are referred to as the "principal components" and are uncorrelated (orthogonal) to each other. The number of principal components generated is equal to the number of original variables. The first principal component accounts for most of the variation in the data, the second principal component accounts for most of the variance that has not been accounted for by the first principal component, and so forth.

Generally, one or two principal components are needed to account for more than half of the variation in the data. As a rule of thumb for project poverty scoring, we suggest using all first principal components necessary to account for at least half of the data variation. Recall that each component is a weighted sum of the variables considered to measure poverty reduction. Thus, higher values for a component denote a higher poverty impact, whereas lower values denote a lower poverty impact.

Steps to Construct a Poverty Scorecard

Given a set of projects that meet a sustainability threshold, the steps to construct a poverty score can be summarized as follows:

1. Collect all the necessary information from the projects to construct the geographic, employment, and spillover indicators (variables) over which the projects will be evaluated.
2. Normalize the indicators by subtracting the mean and dividing by the standard deviation.
3. Obtain the covariance matrix of the normalized indicators.

$$\text{cov}(X, Y) = \sum_{i=1}^n \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}.$$

4. Calculate the eigenvectors and eigenvalues of the covariance matrix. Let A be a $n \times n$ matrix. The parameter λ is an eigenvalue of A if there exists a non-zero vector v such that $Av = \lambda v$. In that case, vector v is called an eigenvector of A corresponding to eigenvalue λ . We can rewrite the condition $Av = \lambda v$ as $(A - \lambda I)v = 0$, where I is the $n \times n$ identity matrix. Now, in order for a non-zero vector v to satisfy this equation, $A - \lambda I$ must not be invertible. That is, the determinant of $A - \lambda I$ must be equal to zero. We call $p(\lambda) = \det(A - \lambda I)$ the characteristic polynomial of A . The eigenvalues of A are simply the roots of the characteristic polynomial of A .
5. Choose the necessary first principal components to account for at least half of the variation in the data. Typically, either the first principal component alone (PC1) or the first two combined (PC1 and PC2) would satisfy this condition.
6. Use the selected principal components in step 5 to rank the projects on the poverty dimension. If more than one principal component is selected, the sum of the two components should be considered for the ranking.

3 An Application to a Grant Competition in Central America

In this section, we apply the scorecard system to a real example of project selection for grant allocation in Central America to link smallholders to markets.

3.1 Details of the Program in Central America

In October 2010, the Unidad Regional de Asistencia Técnica (RUTA), with the technical support of the International Food Policy Research Institute (IFPRI), and with funds from the Multilateral Investment Fund (MIF) of the Inter American Development Bank (IADB) and the Austrian Cooperation in Central America (ADA), launched a program on innovations for linking smallholders to markets in Central America. The nature of the program is pro-poor, market-oriented, and demand-driven, and covers four countries in the region: Guatemala, El Salvador, Honduras, and Nicaragua. In particular, the program is oriented to provide funds (up to a maximum of 250,000 U.S. dollars) for projects that involve creating or improving market opportunities for smallholders in the region, especially in areas with high poverty rates. The projects must be conceived by the same farmers' associations and/or small enterprises applying for the funds.

This program also served as a pilot program to implement the risk and poverty scorecard system described previously. The application process ended in January 2011, the month in which the projects were selected. The projects are currently in the implementation phase.

To ensure transparency, the entire application process was done through a website specifically created for this pilot program. The applicants had to provide

information regarding the association/enterprise applying for the funds, the intended project and its beneficiaries, their credit history (if any), and the project budget and execution plan, as well as any other relevant information. The process also involved filling out some administrative forms.⁷ During the launching of the program, the selection methodology (i.e. the risk and poverty scorecard system) was also explained to the potential applicants.⁸ Related materials and an explanatory video were posted on the website.

A total of 58 projects were submitted by different farmers' associations and small enterprises across the four countries, out of which 39 were eligible for evaluation. The projects that did not qualify were those for which the applicant failed to provide all the required information and forms indicated in the program guidelines. The amount requested by all 39 projects equaled 7.1 million dollars, and a total of 1.7 million dollars was available for grant allocation.

3.2 Risk Scoring

The risk scoring for the 39 projects (applicants) was obtained following the same steps described in Subsection 2.1. Information from previous loans/grants, including performance and characteristics of the borrowers/grantees, was first gathered from other comparable small enterprises in Latin America. These data served to estimate the parameters of interest using the semi-parametric single index model proposed by Klein and Spady (1993). Finally, the estimated parameters, together with the information provided by the 39 applicants, were used to derive their risk scoring.

A total of ten variables were considered to construct the risk scoring. These included the socioeconomic characteristics of both the beneficiaries of the project and the association/enterprise applying for the grant. In particular, we account for average age, gender, education level, and marital status of the beneficiaries and for seniority, number of workers, asset ownership, and financial information of the applicant. For further details on the variables used, see [Table 2](#). The selection of the variables responds to the fact that they are widely used in the microcredit literature, as well as to the amount of previous information available from the pool of loans/grants from similar enterprises. In the pilot program, additional information was naturally collected to validate the information reported by the applicants.

⁷ To apply, applicants had to first register with a username and password. The application did not have to be completed all at once. There were also contact persons in each country for inquiries during the application process. As a proof of their application, once submitted, applicants received a PDF file via e-mail summarizing the information they reported regarding the association/enterprise, project, and direct beneficiaries (basically the information that would then be used to evaluate their application/project).

⁸ The only restriction imposed by the donors for this pilot program was that at least one project had to be selected from each of the four countries. This was also specified at the launching event.

Table 2. List of Variables Used for the Creation of the Risk Scorecard

Characteristics of the direct beneficiaries from the project	Business and financial characteristics of the association / partnership
Average age of beneficiaries	Seniority of association (years of activity)
Gender (percentage of men among beneficiaries)	Number of employees / partners in association
Educational level (percentage of beneficiaries with secondary or superior education)	Assets (if association owns land or real estate)
Marital status (percentage of beneficiaries married or in common-law marriages)	Interest rate of most recent loan
	Term of loan / amount requested

The results from the estimation of the single index model to derive the parameters of interest for the risk scoring are reported in the top panel of [Table 3](#). The sign and magnitude of the coefficients should be interpreted with caution, given that they are normalized with respect to the first variable in the set of explanatory variables (in this case, the average age of the beneficiaries). For matters of comparison and ease of interpretation (although assuming a specific functional and distributional form between the likelihood of defaulting and the explanatory variables), in the bottom panel of [Table 3](#), we also report the results from a standard Probit model. As can be seen, the coefficients generally have the expected sign in the parametric set-up. For example, asset ownership and a lower loan term decrease the probability of default, while older, less educated people and men have a higher probability of defaulting, though the coefficients are not always statistically significant at a 5 percent level in this parametric model.

To show the advantages of using a semi-parametric single index model over a standard Probit model, [Table 4](#) compares the predictive performance of both models. To obtain the predictive performance of each model, the estimated default probabilities (which range between 0 and 1) are first converted to a binary regime prediction using the standard 0.5 rule: if the estimated probability is greater than 0.5, the borrower (grantee) is predicted to default, while if the estimated probability is less than or equal to 0.5, the borrower is predicted to non-default. The binary (1/0) estimated probabilities are then compared to the actual default/non-default behavior of the borrower. It follows that the Probit model has a very poor predictive performance for the default cases (28 percent accuracy versus 86 percent for the single index model). Overall, the predictive performance of the single index model is 72 percent versus 48 percent for the Probit model. These results clearly show the advantages of using a semi-parametric technique over a parametric method for adequately estimating the risk score of a pool of potential borrowers/grantees.

Table 3. Modeling the probability of default (dependent variable equal to one if borrower/grantee defaulted, zero otherwise)

A. Single Index Model		
Variable	Coeff	Std. Err.
Characteristics of beneficiaries		
Average age	1.000	
Percent men	0.136	0.024
Percent with secondary or superior education	−0.250	0.032
Percent married or in common law	0.187	0.025
Characteristics of association		
Years of activity	0.000	0.001
Size (number of workers)	0.016	0.006
If owns land or real state	0.022	0.024
Amount of loan requested (000 dollars)	0.009	0.002
Interest rate previous loan (percent)	0.158	0.019
Term of loan (months)	0.032	0.003
Regression type: local constant		
Model type: Klein and Spady		
Continuous kernel type: Gaussian second order		
Bandwidth:		0.119
# observations		2,899
B. Probit Model		
Variable	Coeff	Std. Err.
Characteristics of beneficiaries		
Average age	0.008	0.004
Percent men	0.115	0.115
Percent with secondary or superior education	−0.136	0.195
Percent married or in common law	−0.111	0.121
Characteristics of association		
Years of activity	−0.005	0.003
Size (number of workers)	−0.019	0.027
If owns land or real state	−0.284	0.116
Amount of loan requested (000 dollars)	0.012	0.014
Interest rate previous loan (percent)	0.002	0.236
Term of loan (months)	0.046	0.017
Constant	−0.056	0.966
# observations		2,899

Table 4. Predictive Performance of Single Index Model and Probit model

Performance indicator	Single index model	Probit model
Overall correct default/non-default classification	71.8 percent	47.8 percent
Correct default classification rate (sensitivity)	86.2 percent	27.6 percent
Correct non-default classification rate (specificity)	51.9 percent	75.7 percent

After applying the estimated coefficients to the information provided by the 39 applicants, we calculated their risk scoring using equation (3). Imposing a cutoff of 67 percent on the maximum risk score allowed, 24 projects qualified for the second stage of the selection process. Please remember that the idea of this first stage in the evaluation process is to identify those projects with a low or moderate level of risk that also ensures the sustainability of the project. The threshold of 67 percent further guaranteed that at least a certain number of projects would get past the risk scoring stage (see [Figure 4](#)). Note also that the projects are not ranked at this stage.

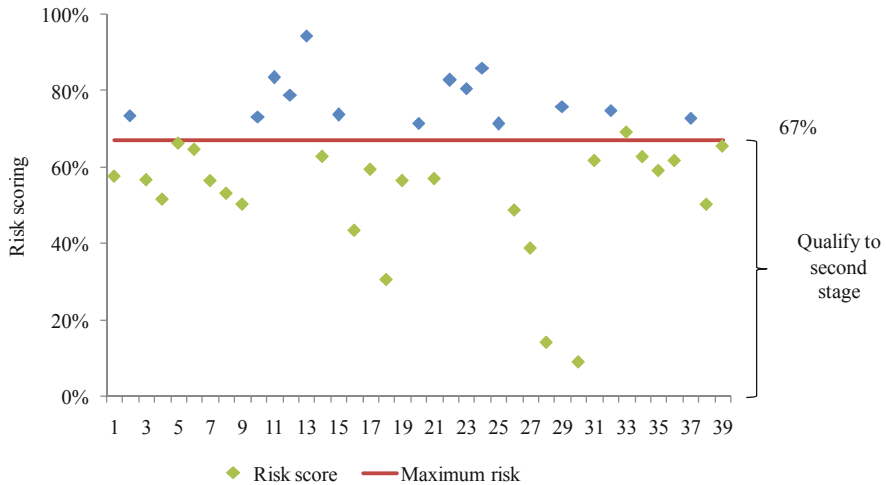


Fig. 4. Risk Scoring Results

3.3 Poverty Scoring

The 24 projects that met the cutoff of 67 percent for the risk scoring were then evaluated in terms of their potential poverty impact following the same steps described in Subsection 2.2. Seven indicators covering the project’s geographic,

Table 5. List of Project Indicators for the Creation of the Poverty Scorecard

Indicator	Variable
Geographic indicators	
1. Location in poverty map-ping space.	Average poverty rate in area(s) where project will be im-plemented.
2. Access to markets	Average distance, in hours, of area(s) where project will be implemented to closest major town with a population of 20,000 or above.
Employment indicators	
3.Labor intensity	Total number of new jobs generated by the project.
4. Low-skill labor intensity	Percent of low-skill labor / total labor in the project.
5. Female labor intensity	Percent of female labor / total labor in the project.
Spillover indicators	
6. Effects on the supply chain	Number of direct beneficiaries / total amount invested.
7. Indirect effects	Number of indirect beneficiaries / total amount invested.

Table 6. Principal Component Analysis: Eigenvalues and Eigenvectors

A. Eigenvalues							
	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Eigenvalue	2.22	1.40	1.11	1.05	0.77	0.45	0.00
Variability (percent)	31.70	19.98	15.84	15.06	10.94	6.47	0.01
Cumulative percent	31.70	51.68	67.52	82.58	93.51	99.99	100.00
B. Eigenvectors							
	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Number of jobs created	0.659	-0.023	-0.169	-0.043	-0.011	0.031	0.730
Percent Low skilled labor	0.014	0.286	-0.234	0.753	0.543	-0.023	-0.004
Percent Female labor	-0.139	0.580	-0.391	-0.390	0.095	0.574	0.007
Direct beneficiaries / amount invested	0.460	-0.317	0.273	-0.157	0.506	0.431	-0.383
Indirect beneficiaries / Amount invested	0.536	0.184	-0.411	0.042	-0.360	-0.241	-0.566
Average poverty rate	0.163	0.370	0.584	0.374	-0.450	0.392	-0.002
Average distance to major closest town	0.141	0.557	0.424	-0.334	0.328	-0.521	-0.005

employment, and spillover characteristics were considered for the poverty scoring and were built based on the information provided by the applicants. The indicators include the average poverty rate and distance to closest major town of the area(s) of project implementation, total number of jobs created, low-skill and female-to-total labor ratio, and number of direct and indirect (intended) beneficiaries per dollar invested (see Table 5). As in the case of risk scoring, additional information was requested in order to validate the information used to construct the indicators.

Poverty scores were then obtained by applying a Principal Component analysis over these seven indicators. The corresponding eigenvalues and eigenvectors derived from the analysis are reported in Table 6. Recall that associated with each eigenvalue, there is an eigenvector and that the total number of principal components generated is equal to the number of variables (indicators) considered for the analysis (in this case, seven). For the poverty scoring and subsequent ranking of projects, we selected the first two components (CP1 and CP2), given that together they account for more than half of the entire variation in the data (52 percent). The first principal component (CP1) is highly correlated with the total number of jobs created and the number of direct and indirect beneficiaries per dollar invested in the project. The second principal component (CP2) is highly correlated with the female-to-total labor ratio, poverty rate, and distance of area(s) of implementation to closest major town.

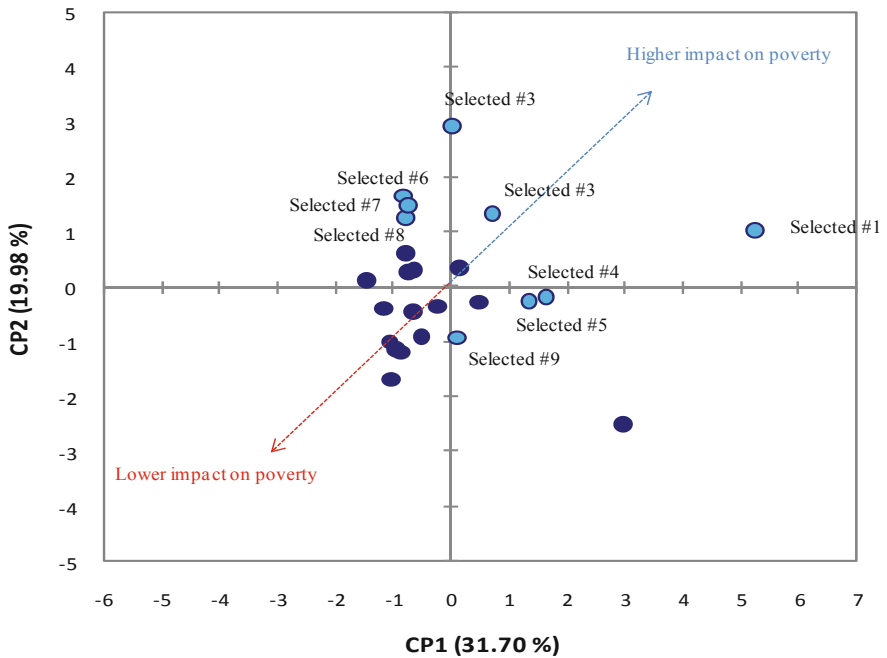


Fig. 5. Poverty Scoring Results

Figure 5 plots the combined scoring of the 24 projects, where the horizontal axis measures the project scoring based on the first principal component and the vertical axis measures the project scoring based on the second principal component. Moving northeast along the 45-degree line implies a higher poverty impact, and vice versa if we move southwest along this line. A total of nine projects were finally selected with a solicited investment of around 1.66 million dollars, slightly less than the 1.7 million dollars allocated for the pilot program. The first eight projects selected were those that obtained the highest poverty score from the Principal Component analysis. The ninth project selected (from Guatemala) strictly obeyed the specific requirement that at least one project from each country be covered by the program.

Table 7 summarizes the distribution of projects (and the amount of funds requested) by country during each stage of the evaluation process: eligible, fulfilled

Table 7. Distribution of Projects Eligible for Evaluation That Get Past the Risk Scoring and That Were Finally Selected by Country

A. Eligible projects		
Country	# Projects	Amount (US\$)
El Salvador	4	786,623
Guatemala	11	1,905,634
Honduras	14	2,697,193
Nicaragua	10	1,725,340
Total	39	7,114,790
B. Projects with a risk scoring below or equal to 67 percent		
Country	# Projects	Amount (US\$)
El Salvador	3	586,713
Guatemala	6	968,946
Honduras	10	1,958,293
Nicaragua	5	955,448
Total	24	4,469,400
C. Selected projects		
Country	# Projects	Amount (US\$)
El Salvador	1	187,075
Guatemala	1	186,700
Honduras	5	992,927
Nicaragua	2	299,759
Total	9	1,666,461

the risk score requirement of 67 percent or less, and selected. As mentioned above, the projects are currently in the implementation phase; a posterior impact evaluation of the selected projects will serve to extend the working database for risk scoring and to improve the scorecard system. The selected projects are intended to benefit nearly 6,000 people (50 percent of which are women) in areas with high poverty levels.

4 Concluding Remarks

This chapter develops a poverty-sensitive scorecard system for lending or grant allocation in developing markets. The proposed methodology used to evaluate loan or grant applications involves constructing both risk and poverty scoring to assure both the sustainability of the project and a high poverty impact. In the first stage, a risk scorecard is developed using a semi-parametric econometric model to assess the default probability of the potential borrowers/grantees. In a second stage, following a Principal Components analysis, a poverty scorecard is constructed to evaluate those projects (borrowers) with a risk score below a certain threshold in terms of their potential poverty reduction. In other words, projects that were proven to be sustainable in the first stage are ranked in the second stage in terms of their potential impact on poverty. We finally implement the scorecard system in a pilot program of grant competition in four countries in Central America to link rural smallholders to markets.

The improvement of the proposed scorecard system over the existing methodology is twofold. First, the use of the latest developments in econometric modeling for the risk scoring allows us to improve the accuracy of risk ranking and to better approximate the sustainability of potential projects under evaluation. Second, by combining two scorecards, our proposed methodology goes beyond poverty scorecards to prioritize lending or grant allocation. Schreiner (2010), for example, recommends a poverty-targeting approach based on identifying high poverty areas using household surveys to prioritize lending. Our proposed lending/grant allocation criterion goes beyond this: it ranks enterprises (projects) with a poverty reduction potential that are also proven to be sustainable. Sustainability is a necessary condition for poverty reduction and a relevant matter in underdeveloped countries, where the problem of adverse selection is acute. Ultimately, the idea is to help policymakers choose from loan or grant applications based on both the chances of project survival and the poverty reduction potential. Additionally, to the extent that we suggest using spatial data and relate to poverty maps, Schreiner's proposed scorecard is nested in the second stage of our scorecard.

The proposed scorecard system should also be viewed as a dynamic process that can be further improved across time, particularly when there is initially limited information for risk assessment. Initial pilot programs, such as the one described in this chapter, could help to provide additional information (based on the performance evaluation of the selected projects) to extend the working database for scoring as-

assessment and to keep improving the scorecard system. This process can continue until a performance history is well established for a reasonably large portfolio of loans/grants. In the long run, the risk scoring mechanism could operate through a simple, implementable program in open-source software (such as a spreadsheet).

Future research should involve formally evaluating the effectiveness of the proposed scorecard system over simple poverty-based targeting and/or over standard risk-based criteria. The further application and extension of pilot programs in developing areas on different contexts will help to perform this exercise.

References

- Adams, D.W., Vogel, R.C. (1985) Rural financial markets in low-income countries: Recent controversies and lessons. *World Development* 14(4):477–487.
- Adams, D.W., Graham, D.H., von Pischke, J.D. (eds.) (1984) *Undermining rural development with cheap credit*. Boulder CO: Westview Press.
- Armendariz de Aghion, B., Morduch, J. (2005) *The Economics of Microfinance*. MIT Press.
- Braverman, A., Guasch, J.L. (1986) Rural credit markets and institutions in developing countries: Lessons for policy analysis from practice and modern theory. *World Development* 14(10/11):1253–1267.
- Brett, C. (2006) Microfinance in Northeast Thailand: Who Benefits and How Much? *World Development* 34(9):1612–1638.
- Ghatak, M., Guinnane, T.W. (1999) The economics of lending with joint liability: theory and practice. *Journal of Development Economics* 60:195–228.
- Ghosh, P., Mookherjee, D., Ray, D. (2000) Credit Rationing in Developing Countries: An Overview of the Theory. In: *A Reader in Development Economics*. London: Blackwell.
- Klein, R.W., Spady, R.H. (1993) An Efficient Semiparametric Estimator for Binary Response Models. *Econometrica* 61:387–421.
- Li, Q., Racine, J. (2006) *Nonparametric Econometrics: Theory and Practice*. Princeton: Princeton University Press.
- McPherson, M.A. (1996) Growth of micro and small enterprises in southern Africa. *Journal of Development Economics* 48(2):253–277.
- Khandker, S. (2005) Microfinance and Poverty: Evidence using panel data from Bangladesh. *The World Bank Economic Review* 19(2):263–286.
- Schreiner, M. (2010) Simple Poverty Scorecards. www.microfinance.com.